Outputs or Outcomes? Assessing Public Relations Evaluation Practices In Award-Winning PR Campaigns

Maureen Schriner, Ph.D.*
University of Wisconsin-Eau Claire

Rebecca Swenson, Ph.D.
University of Minnesota – Twin Cities

Nathan Gilkerson, Ph.D.
Marquette University

For all correspondence:
Maureen Schriner, Ph.D.
Tel: 651 387 3539
e-mail address: schriner@uwec.edu

Abstract

This study advances a framework of concepts based on the Barcelona Principles (AMEC, 2010) to assess the extent public relations practitioners, in award-winning campaigns, place greater emphasis on outcomes than outputs in evaluations of their campaigns, as the Barcelona Principles advocate. The study employs content analysis of the evaluation sections of campaigns recognized for excellence by the Public Relations Society of America Silver Anvil awards from 2010, when the Barcelona Principles were adopted, to 2014. The findings show mixed results. Many cases do measure outcomes, yet the most predominant evaluation described is outputs, which are nearly universally present in campaigns. Awards in certain categories of PR practices are more likely to measure outputs only, with no measurable outcomes. Advertising Value Equivalencies (AVEs), though measured only in a small percent of cases, are still making their way into some evaluation sections. The findings, and recommendations based on the results, provide direction for professional and pedagogical approaches to PR measurement and evaluation.

Keywords: measurement, evaluation, outputs, outtakes, outcomes, AVEs

Introduction

Measurement and evaluation have been an evolving component of strategic communications. Academic and professional organizations alike have placed increasing emphasis on measurement and evaluation in an effort to transition the public relations profession away from an intuitive art form toward greater reliance on scientific methodologies that are both valid and reliable (AMEC, 2015; IPR, 2015). These efforts have coalesced with standards put forth in the Barcelona Declaration of Measurement Principles, otherwise known as the Barcelona Principles (AMEC, 2010). The principles were a consensus of an international coalition of
practitioners and academics from leading organizations in the public relations field to standardize measurement and evaluation with more data-driven, scientific methods (Grupp, 2010; IPR, 2015; Jain, 2014). Five years after the initial declaration, in September 2015, the Barcelona Principles were revised to encompass “holistic” communications, including earned and paid media, as well as owned and shared media channels such as social media (AMEC, 2015).

The purpose of the current study is to expand the framework of the Barcelona Principles by examining measurement and evaluation from the practitioner perspective, focusing on evaluations described in campaigns recognized nationally for excellence through the PRSA Silver Anvil awards. The analysis identifies trends among practitioners, assesses strengths and weaknesses in how best practices in award-winning campaigns are reinforced with the field, and provides direction for measurement and evaluation of strategic communications. To that end, the study first describes the standards, with definitions of key concepts in the Barcelona Principles, to establish a framework for analysis. The Barcelona Principles framework is then used to assess campaign evaluations for Silver Anvil campaigns from 2010 to 2014.

**Literature Review**

2.1 **Historical context for the Barcelona Principles**

The consensus that produced the Barcelona Principles in 2010 follows a long journey of public relations practices, from ancient history to contemporary times (Lamme & Russell, 2009; Lattimore, Baskin, Heiman, & Roth, 2012), a history that has generated successes to the extent that the practice of public relations today is recognized, and applied, at all levels of society around the globe. Throughout the history of public relations, a number of seminal events marked advances in measuring the effectiveness of public relations (Watson, 2012; Likely & Watson, 2013; Volk, 2016). The adoption of the Barcelona Principles is noted by Watson (2012) and others (Grupp, 2010; Thorson, et al., 2015) as a turning point. In putting forth a standardized guide for PR measurement, the individuals and organizations involved in developing the Barcelona Principles recognized two primary problems with measurement and evaluation over the history of public relations. The first problem has been the tradition of PR practices to measure outputs, aspects of PR processes such as counting media clips and measuring the audience reach of news stories about an organization, rather than outcomes, measuring changes in attitude or behavior among publics with whom the organization is communicating, and the effectiveness of PR to advance the organization’s goals (Lindenmann, 2006; Macnamara, 2007; Michaelson & Macleod, 2007).

The second problem has been the lack of scientific, reliable and valid measures in public relations that allow for comparative analysis within the organization, across the PR industry and with other industries (Jain, 2014; Michaelson & Stacks, 2011; White, 2015). Instead, PR practitioners who do measure frequently rely on models developed for internal use by the organization, which are not shared or compared externally, while other PR practitioners simply do not conduct measurements (Thorson, et al., 2015). While PR measurement advocates have stated in-house measurement is better than not measuring at all, they note internal assessments have issues of credibility and possible bias (Lindenmann, 2006).

Advocates for more valid and reliable PR measurements have developed a litany of arguments to advance their cause. Primary among those reasons: to establish value for PR activities in supporting the organization, to be accountable to the organization in justifying the investment in strategic communication, and to use measurement and evaluation to improve the...
performance of PR efforts moving forward (Wilcox & Cameron, 2011; Paine, 2011). Another argument is the need to build respect for public relations as a professional practice within organizations and beyond the PR industry (Rockland, 2010; Zerfass & Sherzada, 2015). Quantifiable outcomes would move PR practices away from what have sometimes been labeled “silly” measurements in public relations (Manning & Rockland, 2011), such as media impressions, advertising value equivalency, and similar methods used to measure PR processes that are not directly tied to organizational outcomes. Without more robust measurements, PR measurement advocates have warned that public relations practitioners will be confined to roles as technicians, or tacticians, rather than strategists in leadership roles of planning, policy-making, advising and management (Macnamara, 2007; Thorson et al., 2015).

Rockland (2014) is more critical of the PR industry’s lackluster use of more rigorous quantitative measurement, stating, “Public relations doesn’t have a measurement problem. Instead, it has an ‘unwillingness to measure’ problem.” Rockland attributes the “unwillingness to measure” problem to a range of weaknesses, from PR practitioners who have poor math skills or lack of knowledge about data-driven methods, to fears of being held accountable for lack of measurable outcomes.

2.2 The Barcelona Principles framework for measurement

As a framework for standardization, the Barcelona Principles offer broad guidelines, with preferred methods, rather than required measurements. The first principle addresses lack of measurement in public relations: “Goal setting and measurement are fundamental to communication and public relations” (AMEC, 2015). In application, the first principle assumes public relations activities are implemented under a strategic plan, in contrast to PR activities based on “seat-of-the-pants” intuition, which van Ruler (2004) conceptualized as a defective, yet common, PR practice. The principles further advocate that “when possible,” PR activities should be measured in terms of their effect on business results. This emphasis on PR activities related to organizational goals and objectives does not directly address public relations’ societal influences, positive or negative, nor the ethics of the organization per se and the organization’s effect on society (Hon, 1997).

The central tenet of the Barcelona Principles, and the focus of the current study, is the standard to measure outcomes as “preferred” to measuring outputs alone. The dichotomy of outputs and outcomes is characterized by key concepts within the current study. In addition, the study discusses outtakes, within the umbrella of outputs applied to the Barcelona Principles, as having distinct attributes that move along the continuum toward outcomes (Lindenmann, 2006).

To apply a consistent, and publicly accessible, framework for the Barcelona Principles, our research uses definitions of measurement levels drawn from one source of collective expertise: the Institute for Public Relations. IPR is one of five organizations listed as authors of the Barcelona Principles (AMEC, 2010; AMEC, 2015). In its role as a resource for the public relations field, IPR provides public online access to two dictionaries. Guidelines for PR evaluation, published by Lindenmann in 2003, and a revision by Stacks and Bowens (2013), provide the field with a compendium dictionary of recently updated concepts for PR measurement and evaluation.

The conceptual framework of outputs, outtakes and outcomes is somewhat controversial in lacking consensus among PR practitioners in terms of how the concepts are defined, or even the necessity for standardized measurements within categories such as outputs, outtakes and
outcomes (Michaelson & Stacks, 2011). The current study addresses the controversy by reinforcing IPR’s role as a centralized, publicly accessible source of standard definitions.

2.3 Differentiating outputs, outtakes and outcomes

The concepts of outputs and outcomes are presented in the Barcelona Principles as contrasting ends of the spectrum. Outputs are processes to distribute messages, or “the number of communication products or services resulting from a communication production process” (Stacks & Bowen, 2013). The measurements are “the immediate results of a particular PR program or activity,” (Lindenmann, 2003) which measure potential public exposure for the organization. Examples of output measurements include the number of brochures distributed publicly, the amount of online communications produced, the number of events held and the number of participants at events, counts of press releases and other contacts with news media, and volume of news coverage (media clips). Outputs have traditionally had an assumed value and an assumed impact on audiences.

At the other end of the PR measurement spectrum, outcomes measure change in the targeted audience’s cognition, affect and behavior (Lindenmann, 2003). The outcomes are to support the organization’s overall goals and objectives, as opposed to arbitrary objectives set specifically for the PR campaign. Stacks and Bowen (2013) list five levels of change to measure: awareness, knowledge, attitude, opinion, and behavior levels. The critical characteristic of measurement in this case is “change,” which is validated through pre-test and post-test measurements. The Barcelona Principles in 2010 establish a “preferred” method of measuring change through benchmarking and tracking.

Outtakes offer a midpoint along the continuum between output and outcome. Within the PR standards of the Barcelona Principles, outtakes fall under the umbrella of outputs, as they do not reach the standards of outcomes. Outtakes measure if the intended audience actually received, and cognitively processed, the PR message (Lindenmann, 2003). Outtakes are “what audiences have understood and/or heeded and/or responded to a communication product’s call to seek further information from PR messages prior to measuring an outcome” (Stacks & Bowen, 2013). In other words, the targeted audience is taking steps to interact with the message sender, which is an important point in establishing relationships (Hon & Grunig, 1999). Outtakes have four units of measure (Lindenmann, 2003): favorability, such as whether news media reported stories in positive, negative or neutral tones; comprehension, determining how recipients interpreted the message; recall and retention, the extent an embedded message is memorable to the receivers; and favorable responses, such as whether the receiver sought more information on a website, “liked” a Facebook post, or shared the information with others. While outtakes move toward outcomes, the Barcelona Principles subsume outtakes within outputs, as both are steps in strategic communication processes.

Lindenmann (2003) notes that measuring outcomes is “usually much more difficult and, generally, more expensive” than measuring outputs or even outtakes, but measuring outcomes is “far more important” than the easier route of measuring outputs and outtakes.

2.4 Evaluating news media

The Barcelona Principles call for media to be measured for quantity and quality. Evaluating media is important given the dominance of media relations as a PR practice (Jamison, Gaines-Ross, & Gorman, 2014). So-called “earned media” initiatives, or campaigns focused on gaining media placements, remain a key element of the practice. Quantity of media is measured by volume of news coverage and media impressions, which are categorized as outputs (Stacks &
Bowen, 2013). Quality of media involves measuring news content for tone, placement, visuals, whether the news story carried key messages and contained third-party endorsements, as well as assessing the credibility and audience reach of the media outlet (Grupp, 2010). Quality analysis of news media is an outtake by measuring response from an intermediary audience (Stacks & Bowen, 2013).

Media measurements would reach the level of outcome if the evaluations were to benchmark knowledge and attitudes pre- and post-news coverage, to determine the news stories’ influences on the targeted audience (Michaelson & Macleod, 2007).

2.5 Advertising value equivalents

Advertising value equivalents, or AVEs, compare earned media to paid media, through a process of assigning an advertising “value” to news coverage. The third principle of the Barcelona Principles unequivocally rejects AVEs in stating, “AVEs are not the value of communications” (AMEC, 2015). Stacks and Bowen (2013) describe AVEs as “a discredited output score.” As Jeffries-Fox (2003) explains, there are significant conceptual differences between advertising and news media coverage that negate the value placed in the AVE comparison.

2.6 Measuring social media

As PR practitioners expand their use of social media for distributing messages (Wright & Hinson, 2014), measurement and evaluation of social media has developed challenges, such as how PR efforts in social media demonstrate value to the organization and support organizational goals and objectives (McCorkindale & DiStaso, 2014). The Barcelona Principles call for PR practitioners to evaluate social media along the same lines as evaluation of news media coverage in quality and quantity. Invoking the PR industry’s interest in engagement, the principles state, “Measurement must focus on ‘conversation’ and ‘communities,’ not just ‘coverage’” (AMEC, 2015). Website visits, Facebook “likes” or Twitter tweeting activities, along with a number of other web analytics and social media measures, are categorized as either outputs or outtakes by Stacks and Bowen (2013), depending on whether the measurements show the receiving public is interacting with the PR messenger. To achieve the level of measuring outcomes, the Barcelona Principles recognize the need for technology-assisted analysis to measure social media, particularly given the volume and scope of communications generated by social media (Grupp, 2010). Public relations practitioners face skepticism among business executives in showing how social media connects to business outcomes and influences key stakeholders (DiStaso, McCorkindale, & Wright, 2011), which reinforces the need for data analysis to demonstrate social media interactions have moved beyond PR processes to change attitudes or behaviors in support of the organization.

2.7 Evaluating events

Events, similar to media relations, pose a measurement challenge, in that it is often difficult to measure whether messages from events reached the intended targeted public and even more challenging to determine how the events may have influenced targeted publics to support organizational goals and objectives. Another challenging factor is the generally limited scope of participants in most special events.

Hardy and Waters (2012) raised concern about the predominance of special events in their analysis of 42 years of Silver Anvil campaigns. They warn that overreliance by PR practitioners on special events and media blitz campaigns could indicate public relations is more tactical than strategic, and it has not fully shifted to become a management function.
In sum, the framework for the Barcelona Principles calls for public relations practitioners to transition from PR processes and the counting of outputs, to focus on evaluation of outcomes, with rigorous measurements of changes in attitudes and behaviors, which support organizational goals and objectives. The framework is applied in the current study in a specific context, examining evaluations in campaigns recognized for their excellence by the U.S.-based Public Relations Society of America (PRSA). The Silver Anvil awards are self-described by PRSA (2015) as “the pinnacle of achievement and the highest standards of performance in the public relations profession.” The awards recognize campaigns conducted by organizations in all sectors, from campaigns conducted locally to globally. The framework applied by Hardy and Waters (2012) in analyzing Silver Anvil awards over a 42-year period, was to identify use of elements in the ROPES strategic planning process. The study found formative research had increased in volume and range of methods, and that campaigns had shifted from measuring awareness and attitudes to greater emphasis on measuring actions, with the exception of evaluations for special events.

In the years covering the study, PRSA Silver Anvil award entries are summarized in template forms with sections that cover each step of the strategic planning process. As such, award applicants are expected to cover each section, including the final section on evaluation. The template, with its evaluation section, provides a standardized data set to explore PR practices in measurement and evaluation. The current study focuses on the evaluation sections as follows:

- **R1:** To what extent do Silver Anvil cases provide outputs or outcomes in their evaluations?
- **R2:** What trends emerge in Silver Anvil cases in evaluation practices by PR practitioners between 2010 and 2014 in measuring outputs and outcomes?

The intent of the research questions is to reveal characteristics of public relations campaigns at the highest levels of the PR field as an indication of how measurement and evaluation are applied in practice and how best practices are reinforced within the profession.

**Method**

Analysis in this study focused on Silver Anvil cases over a 5-year period, from 2010, when the Barcelona Principles were declared, to 2014. The cases, posted on the PRSA Silver Anvil archives, are available to PRSA members on the PRSA website. Cases from 2010 to 2014 were collected for analysis by downloading the PDF file of each case. The analysis included cases that won the Silver Anvil (first place) and cases listed as Awards of Excellence (runner-ups). Although the award titles differ in the award ceremony itself, the first place and runner-up awards are equally represented on PRSA’s website as award-winning excellence in PR campaigns. Content for the analysis was drawn solely from the case as described in the PDF file and did not address details in the full application submitted to PRSA. The case summary is presumed to be an accurate reflection of the case as a whole, given that the PDF file in summary form is how the case is presented to PRSA members, and to the world.

A total of 526 cases were included in final coding. Within any given year, cases awarded in more than one category were analyzed only once, to avoid having evaluations from one campaign carry more influence than other campaigns. Cases in which the evaluation section was not completed or not clearly labeled as the evaluation section were excluded.

The coding narrative was intended to reflect the intentions of the PR practitioners to the extent they placed importance on each evaluation they described. The coding assessed content from the evaluations and coded for the dichotomous variable, output or outcome, for each
measurement included in the evaluation section. As such, the content from one campaign could include in its evaluation numerous outputs and outcomes. Advertising value equivalencies were coded in cases in which the evaluation measure directly mentioned advertising values or AVEs, indicated by dollar signs for the values, which included public service announcements that tied the value directly to advertising value along with dollar sign values. For intercoder reliability, two coders analyzed evaluation sections for 60 cases (11 percent of total cases), with a random sampling of 12 cases from each year (2010-2014). Cohen’s Kappa reached .75 in distinguishing between outputs and outcomes.

The methodology was descriptive analysis, assessing both quantitative and qualitative factors. Due to the exploratory nature of the research questions and the limited focus of the dataset, the study did not use statistical analysis.

Results

Regarding RQ1, the findings support the efforts of the Barcelona Principles in advocating for PR campaigns to evaluate outcomes. As summarized in Table 1, which provides counts of cases coded, the cases with outcomes in the evaluation sections remain consistently high over the five years. The award winners’ emphasis on outcomes appears to reflect the level of excellence expected for PR campaigns nationally recognized for excellence.

Table 1
Total cases coded and cases with outcomes

At the same time, the award-winning campaigns place even greater emphasis on output measures within their campaign evaluations. As Table 2 illustrates, the percent of campaigns to include outputs as evaluations was nearly universal, exceeding the percent of campaigns that included outcome evaluations. There were few campaigns that evaluated only outcomes, as the practice in the award-winning campaigns was to report outputs along with outcomes. Often, the case presentation implied causality, with outputs ostensibly resulting in outcomes. In campaigns
with no outcome evaluations, the case implied some type of positive result from the outputs, but did not meet the standards for outcomes, often due to lack of pre-test/post-test measures or benchmarking.

Table 2 Evaluations of cases with percent listing outputs, and percent listing outcomes

Two findings address RQ2, in identifying trends in measurement. One trend is the ongoing use of AVEs. As illustrated in Table 3, the percent of case evaluations reporting AVEs has remained around 10 percent over the 2010-2014 period. While this is a relatively low percent, the persistent use of AVEs, and other descriptions of PR whose value is directly tied to paid advertisement placements, may indicate that PR practices need to focus on further developing alternative measures that provide clear value and are readily understood by other business management functions and decision-makers within an organization. Organizers of the conference that produced the Barcelona Principles noted that while there was universal consensus to reject AVEs as a measurement of the value of public relations, the conference was unable to come to consensus on an alternative to AVEs (AMEC, 2010).

Table 3 Percent cases with AVE evaluation measures
Given the consistency of the percentages of outputs and outcomes in campaign evaluations from 2010 to 2014, the authors chose not to compare differences in campaigns from year to year. Instead, five years of cases were compiled to compare campaigns that did evaluate outcomes to those that did not, based on category of award. During the 5-year period of analysis, there were 13 Silver Anvil categories in which PR practitioners had submissions. Table 4 lists the categories in rank order, starting with the most commonly awarded category.

Table 4
Categories of awards by percentage of total cases

<table>
<thead>
<tr>
<th>Award type</th>
<th>Percentage of total cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marketing consumers</td>
<td>21.7%</td>
</tr>
<tr>
<td>Events and observances</td>
<td>17.9%</td>
</tr>
<tr>
<td>Integrated communications</td>
<td>9.5%</td>
</tr>
<tr>
<td>Reputation management</td>
<td>9.5%</td>
</tr>
<tr>
<td>Public service</td>
<td>8.9%</td>
</tr>
<tr>
<td>Community relations</td>
<td>6.8%</td>
</tr>
<tr>
<td>Marketing B2B</td>
<td>6.3%</td>
</tr>
<tr>
<td>Internal communications</td>
<td>4.9%</td>
</tr>
<tr>
<td>Public affairs</td>
<td>4.2%</td>
</tr>
<tr>
<td>Crisis communications</td>
<td>3.4%</td>
</tr>
<tr>
<td>Multicultural</td>
<td>3.4%</td>
</tr>
<tr>
<td>Issues management</td>
<td>1.9%</td>
</tr>
<tr>
<td>Investor relations</td>
<td>1.50%</td>
</tr>
</tbody>
</table>

In the outcome/ no outcome comparison, categories with few awards did not reveal consistent trends. Issues management and investor relations, which had the fewest award-winning submissions, had all campaigns include outcome evaluations. Campaigns awarded for multicultural communications, also with few submissions, had half the campaigns evaluate outcomes. The outcome/ no outcome comparison did reveal a pattern when focused on the five most popular categories. Table 5 provides insight for RQ2 in identifying the category of events and observances as more likely than any other category to not have outcomes in campaign evaluations. But even the most popular category, marketing to consumers, had about one in four campaigns that did not evaluate outcomes.
In sum, the results indicate the need to further develop methods of measurement that not only direct PR practitioners, but address common PR practices of measuring outputs, especially for particular categories of public relations activities.

Discussion

Overall the findings of content analysis of evaluations in PRSA Silver Anvil campaigns from 2010 to 2014 support the general direction of the Barcelona Principles, as many cases did evaluate outcomes of attitudinal or behavioral change, meeting the high bar set in the Barcelona Principles. The emphasis on quantifiable evaluations by leaders in the PR industry is a means to establish best practices in the field. At the same time, these nationally recognized PR campaigns are presenting a consistent progression in the evaluation process, to emphasize outputs along with outcomes. This trend is worthy of future research and discussion in the public relations field, in understanding measurement and evaluation in the context of PR practices placing high value on outputs in evaluating their success.

The findings of this study provide evidence to support direction in three areas that have already been raised in the literature. The first directional area is for PR practitioners to add value to their activities, whether they are outputs or outcomes. In proposing a “toe-bone-to-head-bone” logic model for measurement and evaluation, Macnamara (2013) advances the idea that
PR practitioners evaluate their work in terms of stepping stones, milestones or “sub-outcomes” to support the organization’s goals and objectives. The logic model allows for outputs to carry value in evaluation in supporting organizational goals. Just as other divisions within an organization are not required to prove every action would directly connect to macro-level outcomes for the organization, the strategic communicators should explain the contributions of outputs and outcomes in steps, making connections from toe to head, as Macnamara explains. “If [communication tactics] are not connected in some way to higher-level outcomes, it is appropriate to question their value.” This stepping-stone approach to measurement is important not only for campaigns to support organizational goals and objectives, but in supporting PR activities in campaigns of broader social nature, in which a single organization does not control the campaign goals and objectives.

A second area for direction in PR measurement is the finding that events and observances were more likely as a category than others to have no measurable outcomes. These findings reinforce the findings of Hardy and Waters (2012), which indicate public relations practices are overly reliant on special events and publicity stunts. In the context of the measurement movement advocates, the PR industry will continue to be viewed as publicity tacticians particularly if activities such as events are not measured for their value to the organization’s goals and objectives. Additionally, as one of the most frequent PR practices, events and observances could benefit from development of specific methods, and industry benchmarking, to connect the value of special events to organizational goals and objectives.

A third directional area is continued emphasis on best practices. As the findings of this study reveal, the PR industry is generating campaigns that evaluate outcomes. Based on the trends identified in this study, best practices emerge when PR practitioners apply rigorous methods for measuring outcomes. This becomes increasingly important for PR in demonstrating to executive management the value of social media and digital engagement with publics. Measuring change in attitudes and behaviors requires pre- and post-testing, and benchmarking, research methodologies that academic educators should emphasize in pedagogical approaches to instruct students learning about strategic communication.

6. **Conclusion**

In summary, the PR standards in measurement and evaluation, developed as an international consensus in the Barcelona Principles, are being practiced in award-winning PR campaigns, with some exceptions. Many cases include outcomes in their evaluations yet continued to place greater emphasis on output measures. Categories for events and observations were more likely to have no outcomes evaluated, pointing to the need to develop effective outcome measurements for events. AVEs continue to have a small, but persistent, presence in evaluations among award-winning campaigns.

**6.1. Limitations**

Some of the limitations of this study have been previously described above. An additional limitation affecting the findings is that PRSA Silver Anvil award-winning campaigns tend to be shorter-term PR initiatives, not long-term programs that are better structured to measure and evaluate outcomes through benchmarking. An additional limitation is that the PRSA case content is dependent on the skills and knowledge of the PR practitioners writing up the award case, which would influence what and how evaluations are recorded.
Acknowledgements
The authors would like to thank the reviewers at the 18th International Public Relations Research Conference, who provided advice and guidance on this research.

References


